

ESERCIZI - Principi di Inferenza

PARTE I : Statistica descrittiva

B. Franzolini, beatrice.franzolini@unimib.it

1. Su un campione di 90 studenti è stato rilevato il numero di errori fatti in un compito ottenendo la distribuzione di frequenze:

Numero di errori	Frequenza
1	36
3	27
4	18
7	9

- a) Determinate le frequenze relative (e percentuali).
- b) Fornite un'opportuna rappresentazione grafica.
- c) Determinate la frequenza relativa con cui gli studenti del campione fanno:
- meno di 3 errori
 - al massimo 3 errori
 - più di 2 errori e al massimo 4
 - almeno 2 errori e meno di 4
2. La tabella seguente riporta dati ISTAT dal censimento 2011

Scuola superiore scelta	Alunni Italiani (in migliaia)	Alunni Stranieri (in migliaia)
Liceo	436	48
Istituto tecnico	168	26
Istituto professionale	270	56
Totale	874	130

- a) Quali sono le due popolazioni considerate?
- b) Qual è la variabile rilevata su entrambe le popolazioni?
- c) Di che tipo di variabile si tratta?
- d) Che cos'è il numero 168 in tabella?
- e) Quali sono le due distribuzioni di frequenze relative?
- f) Quale grafico useresti per rappresentare la distribuzione di frequenze corrispondente agli alunni italiani? Perché?
3. La tabella seguente contiene la distribuzione delle frequenze del carattere età sulla popolazione degli italiani che hanno fatto attività di volontariato nel 2015. Completa la tabella.

Età	Frequenze assolute (in migliaia)	Frequenze percentuali
15-34	1590.4	28.4%
35-54		
55-90		35.2%
Totale		

4. Considera la seguente tabella realizzata a seguito di un'indagine censuaria sui 23 500 studenti iscritti all'università di Perugia.

Numero componenti del nucleo familiare	Frequenza (in migliaia)
2	1.3
3	12.5
4	8.5
5	1.17
6	0.01
7	0.02

- Qual è la popolazione?
 - Qual è la variabile rilevata?
 - Di che tipo di variabile si tratta?
 - Che cos'è il numero 6 in tabella?
 - Qual è la distribuzione di frequenze relative?
 - Quale grafico useresti per rappresentare la distribuzione di frequenze? Perché?
5. Si consideri la distribuzione di frequenza relativa delle famiglie italiane per numero di componenti nel 2017 (Fonte Istat):

Numero componenti	1	2	3	4	5	≥ 6	Totale
Frequenze relative	0.319	0.275	0.196	0.157	0.041	0.012	1.000

- Si determini la distribuzione di frequenza relativa cumulata.
 - Si determini la distribuzione di frequenza assoluta (con le frequenze in migliaia), sapendo che il numero totale delle famiglie era 25 865 (in migliaia).
6. Un sito di e-commerce ha analizzato il **numero di articoli acquistati in un mese** da un campione di 30 clienti registrati al programma fedeltà.

I dati sono stati sintetizzati nella seguente tabella di frequenze:

Numero di articoli	Frequenza assoluta
0	2
1	4
2	6
3	8
4	5
5	3
7	2

- Calcolare la **media aritmetica** del numero di articoli acquistati.
- Calcolare la **mediana**.
- Calcolare la **moda**.
- Qual è la percentuale di clienti che ha acquistato più di 1 articolo durante il mese?
- Calcolare il **primo quartile** $Q_{0.25}$ e il **terzo quartile** $Q_{0.75}$.
- Qual è il quantile-0.80? Come si interpreta questo valore?

7. Dati le seguenti osservazioni di cui la tabella riporta le frequenze assolute:

Valore	2	3	4	5
Frequenza	4	6	3	2

Calcola la media e la distribuzione di frequenze relative.

8. I dati seguenti sono stati raccolti durante un'indagine sull'uso del telefono e lettura di libri su un campione di 12 persone.

Età	# telefoni	Screen-time telefono in ore al giorno	Spesa per libri in euro nell'ultimo mese	Letto almeno un libro nell'ultimo mese?
19	1	3	0	No
16	1	8	30	Si
60	1	2	10	No
24	2	5	0	No
49	3	2	115	Si
15	1	4	25	No
28	2	1	96	Si
47	1	3	30	Si
21	2	5	20	No
21	1	0	108	Si
57	1	4	16	Si
17	1	3	20	No

- Mediamente hanno uno screen-time più elevato le persone che hanno letto un libro nell'ultimo mese o quelle che non lo hanno letto? Motivare la risposta.
- La mediana dello screen-time è più elevata per le persone che hanno letto un libro nell'ultimo mese o per quelle che non lo hanno letto? Motivare la risposta.

9. La tabella che segue riporta i valori dell'attivo di bilancio, del fatturato e del debito finanziario (in migliaia di euro) per un campione di società quotate e non quotate iscritte nel Registro delle imprese.

Nome Impresa	Tipo di società	Attivo di bilancio	Fatturato	Debito Finanziario
Valerio Marziale	Quotata	250	480	15
Tullio Cicerone	Non Quotata	175	235	100
Plinio il Giovane	Quotata	480	100	200
Appio Claudio	Quotata	235	300	165
Cornelio Nepote	Quotata	180	210	205
Severino Boezio	Non Quotata	133	120	85
Lucio Seneca	Non Quotata	95	100	25
Giulio Cesare	Non Quotata	235	200	45

- Si determinino le medie del debito finanziario per le società quotate e non quotate. Quale tipologia di società risulta maggiormente indebitata?
- Si considerino le variabili "fatturato" e "attivo di bilancio" per l'intero campione. Quale carattere presenta la maggiore variabilità?

10. Ad un campione di 200 clienti abituali di un supermercato è stato chiesto di dichiarare il numero di utilizzi, nello scorso mese, del servizio di farmacia annesso al magazzino. I dati sono sintetizzati in tabella:

n. utilizzi	n. clienti
0	98
1	44
2	44
3	11
5	3

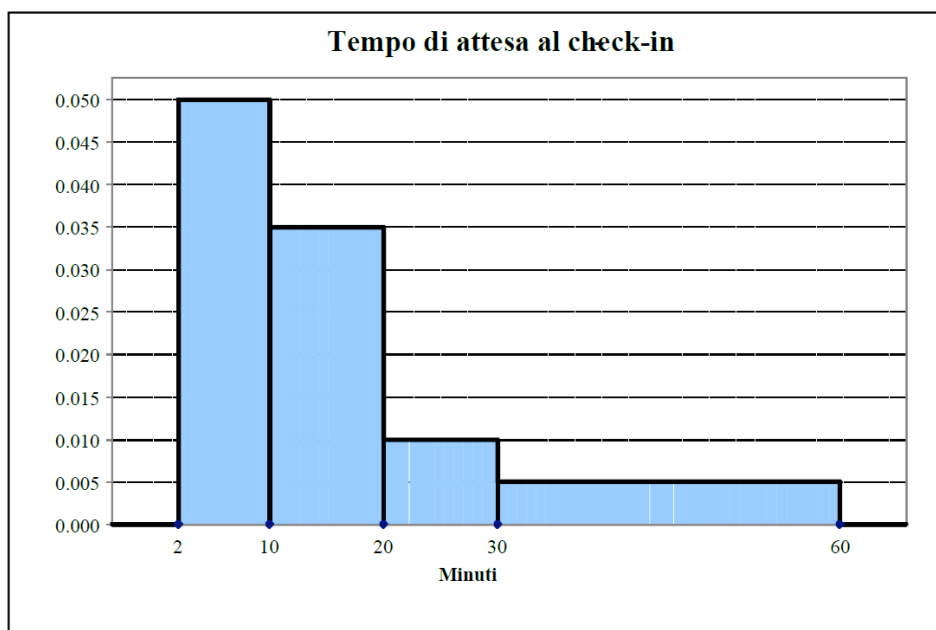
- a) Si specifichi la tipologia della variabile “numero di utilizzi”.
- b) Si calcoli la media aritmetica del numero di utilizzi.
- c) Si calcoli la varianza del numero di utilizzi.
11. Una compagnia di telefonia mobile che opera su tutto il territorio nazionale ha condotto un’indagine sui propri clienti. La seguente tabella riporta i dati di un campione casuale di 9 clienti, titolari di un contratto prepagato, per i quali sono state rilevate le seguenti variabili:
- **MINUTI**: minuti di conversazione al cellulare nell’ultima settimana;
 - **RICARICA**: taglio dell’ultima ricarica effettuata (euro);
 - **SPESA**: spesa sostenuta per l’acquisto del cellulare (euro).

I dati sono riassunti nella tabella che segue:

	MINUTI	RICARICA	SPESA	MINUTI ²	SPESA ²
	38	50	102	1444	10404
	15	25	156	225	24336
	45	80	599	2025	358801
	18	25	299	324	89401
	22	25	365	484	133225
	26	50	178	676	31684
	20	50	214	400	45796
	42	80	188	1764	35344
	40	50	376	1600	141376
Totale	266	435	2477	8942	870367

- a) Si determini la distribuzione di frequenze della variabile **MINUTI** utilizzando 3 classi di uguale ampiezza.
- b) Si determini la mediana della variabile **RICARICA**.
- c) Si determini il coefficiente di variazione della variabile **SPESA**.

12. Il seguente istogramma evidenzia, per un campione di 400 passeggeri di voli intercontinentali, il tempo di attesa al check-in (in minuti). (Sulle ordinate sono riportate le densità di frequenza.)



- a) Si determinino le frequenze relative di ogni classe.
- b) Si determini la percentuale di coloro che attendono tra 5 e 27 minuti.